

Form PTO-1390		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER P21250
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		U.S. APPLICATION NO. (If known, see 37 CFR 1.5) <b>09/869312</b>	
INTERNATIONAL APPLICATION NO. PCT/IP00/00355	INTERNATIONAL FILING DATE 25 January 2000	PRIORITY DATE CLAIMED 25 January 1999	
TITLE OF INVENTION DESCRIBING AND STORING METHOD OF ALIGNMENT INFORMATION			
APPLICANT(S) FOR DO/EO/US Tetsuro TOYODA and Akiko ITAI			
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information.			
<p>1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.</p> <p>2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.</p> <p>3. <input checked="" type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)).</p> <p>4. <input checked="" type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).</p> <p>5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))</p> <p>a. <input checked="" type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau).</p> <p>b. <input checked="" type="checkbox"/> has been communicated by the International Bureau.</p> <p>c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</p> <p>6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).</p> <p>7. <input type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))</p> <p>a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).</p> <p>b. <input type="checkbox"/> have been communicated by the International Bureau.</p> <p>c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</p> <p>d. <input type="checkbox"/> have not been made and will not be made.</p> <p>8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).</p> <p>9. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).</p> <p>10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (U.S.C. 371(c)(5)).</p>			
Items 11 to 16 below concern other document(s) or information included:			
11. Assignee: <u>INSTITUTE OF MEDICINAL MOLECULAR DESIGN, INC. OF Tokyo, JAPAN</u>			
12. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.			
13. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.			
14. <input checked="" type="checkbox"/> A FIRST preliminary amendment.			
<input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.			
15. <input type="checkbox"/> A substitute specification.			
16. <input type="checkbox"/> A change of power of attorney and/or address letter.			
17. <input type="checkbox"/> Figure of Drawing to be published _____			
18. <input checked="" type="checkbox"/> Other items or information:			
Cover Sheet and International Application as published.			
PCT/RO/101-PCT Request.			
PCT/IPEA/416.			
PCT/IPEA/409.			
PCT/IPEA/408.			
PCT/IB/301.			
PCT/IB/304.			
PCT/IB/308.			
PCT/IB/332.			
PCT/ISA/220.			
PCT/ISA/210.			
Cover Letter under 35 USC 371 and 1.495.			
Claim of Priority.			

U.S. APPLICATION NO. (if known, see 37 CFR 1.5) <div style="font-size: 2em; font-weight: bold; margin-top: 5px;">097869312</div>	INTERNATIONAL APPLICATION NO. PCT/JP00/00355	ATTORNEY'S DOCKET NUMBER P21250																										
19. The following fees are submitted:		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 50%;">CALCULATIONS</th> <th style="width: 50%;">PTO USE ONLY</th> </tr> </table>	CALCULATIONS	PTO USE ONLY																								
CALCULATIONS	PTO USE ONLY																											
Basic National Fee (37 CFR 1.492(a)(1)-(5)):  Search report has been prepared by the EPO or JPO. .... \$ 860.00 International preliminary examination fee paid to USPTO (37 CFR 1.482). .... \$ 690.00 No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)). .... \$ 710.00 Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO. .... \$1,000.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4). .... \$ 100.00  <div style="text-align: right;">ENTER APPROPRIATE BASIC FEE AMOUNT =</div>		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$660.00</td> </tr> </table>		\$660.00																								
	\$660.00																											
Surcharge of \$130.00 for furnishing the oath or declaration later than ___ 20 ___ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$</td> </tr> </table>		\$																								
	\$																											
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 15%;">Claims</th> <th style="width: 15%;">Number Filed</th> <th style="width: 15%;">Number Extra</th> <th style="width: 15%;">RATE</th> <th style="width: 15%;"></th> <th style="width: 15%;"></th> </tr> </thead> <tbody> <tr> <td>Total Claims</td> <td>16</td> <td>- 20 =</td> <td>0</td> <td>X \$18.00</td> <td>\$0.00</td> </tr> <tr> <td>Independent Claims</td> <td>4</td> <td>- 3 =</td> <td>1</td> <td>X \$80.00</td> <td>\$80.00</td> </tr> <tr> <td>Multiple dependent claim(s) (if applicable)</td> <td></td> <td></td> <td></td> <td>+ \$270.00</td> <td>\$0.00</td> </tr> </tbody> </table>		Claims	Number Filed	Number Extra	RATE			Total Claims	16	- 20 =	0	X \$18.00	\$0.00	Independent Claims	4	- 3 =	1	X \$80.00	\$80.00	Multiple dependent claim(s) (if applicable)				+ \$270.00	\$0.00	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$940.00</td> </tr> </table>		\$940.00
Claims	Number Filed	Number Extra	RATE																									
Total Claims	16	- 20 =	0	X \$18.00	\$0.00																							
Independent Claims	4	- 3 =	1	X \$80.00	\$80.00																							
Multiple dependent claim(s) (if applicable)				+ \$270.00	\$0.00																							
	\$940.00																											
TOTAL OF ABOVE CALCULATIONS =		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$940.00</td> </tr> </table>		\$940.00																								
	\$940.00																											
Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$</td> </tr> </table>		\$																								
	\$																											
SUBTOTAL =		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$940.00</td> </tr> </table>		\$940.00																								
	\$940.00																											
Processing fee of \$130.00 for furnishing the English translation later than ___ 20 ___ 30 months from the earliest claimed priority date (37 CFR 1.492(f)).		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">+</td> </tr> </table>		+																								
	+																											
Extension of Time fee in the amount of \$		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%;"></td> </tr> </table>																										
TOTAL NATIONAL FEE =		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$940.00</td> </tr> </table>		\$940.00																								
	\$940.00																											
Fee for recording the enclosed assignment (37 CFR 1.21(h). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">+</td> </tr> </table>		+																								
	+																											
TOTAL FEES ENCLOSED =		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$940.00</td> </tr> </table>		\$940.00																								
	\$940.00																											
Amount to be refunded		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$</td> </tr> </table>		\$																								
	\$																											
Charged		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;"></td> <td style="width: 50%; text-align: center;">\$</td> </tr> </table>		\$																								
	\$																											
a. <u>X</u> A check in the amount of \$940.00 to cover the above fees is enclosed. b. ___ Please charge my Deposit Account No. ___ in the amount of \$ ___ to cover the above fees. c. <u>X</u> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>19-0089</u> .  NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.  SEND ALL CORRESPONDENCE TO CUSTOMER NO. 7055 AT THE PRESENT ADDRESS OF: Bruce H. Bernstein GREENBLUM & BERNSTEIN, P.L.C. 1941 Roland Clarke Place Reston, VA 20191 (703) 716-1191																												
<div style="text-align: right;">           SIGNATURE          Bruce H. Bernstein          NAME          29,027          REGISTRATION NUMBER       </div>																												

P21250.A01

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Tetsuro TOYODA et al.

Serial No : Not Yet Assigned (National Stage of PCT/JP00/00355)

Filed : Concurrently Herewith (International Filing Date January 25, 2000)

For : DESCRIBING AND STORING METHOD OF  
ALIGNMENT INFORMATION**PRELIMINARY AMENDMENT**Commissioner of Patents and Trademarks  
Washington, D.C. 20231

Sir:

Prior to calculation of the filing fees and the examination of the above-identified patent application on the merits, the Examiner is respectfully requested to amend the claims as follows:

IN THE CLAIMS

Please amend the claims as follows (a marked-up copy of the claim amendments is provided as an attachment to this Amendment):

5. (Amended-Clean Text) The method according to claim 1, in which the gap information is described by using the residue number and/or the number of residues which indicates the location and the length of a gap region existing in the alignment information of more than two sequences, or a numerical data convertible to those numbers by calculation.

P21250.A01

6. (Amended-Clean Text) The method according to claim 1, in which the gap information is described by using a data containing the residue number of another sequence not included in the alignment or a virtual sequence or column number of a standard representation of the alignment information or a numerical data convertible to the data.

7. (Amended-Clean Text) A method of generating a gap information of a new alignment information by calculating on only a gap information generated from an information of one or more alignments by any one of the methods according to claim 1.

8. (Amended-Clean Text) A method of obtaining a standard representation of an alignment information from both a gap information and a sequence information obtained by the method according to claim 1.

12. (Amended-Clean Text) A data record including at least its identifier and a gap information and identifiers of sequences in a sequence information generated by the method according to claim 1, and a media containing one or more of the record.

P21250.A01

13. (Amended-Clean Text) The data record according to claim 12, wherein its identifier is generated by a method of determining a substantially unique identifier of an alignment information, wherein the identifier is determined using and depending on only and all of a data on a gap information and identifiers of sequences in a sequence information.

14. (Amended-Clean Text) A method of storing, searching and/or communicating method of an alignment information which employs at least the data record according to claim 12.

15. (Amended-Clean Text) The method according to claim 1 which employs the correspondence information instead of the gap information.

16. (Amended-Clean Text) The method according to claim 1 which employs eigen-identifier of sequences in the sequence information.


#### REMARKS

By the above amendment, the claims have been amended to delete multiple dependency.

P21250.A01

If there should be any questions, the Examiner is invited to contact the undersigned  
at the telephone number listed below.

Respectfully submitted,  
Tetsuro TOYODA et al.

  
Bruce H. Bernstein  
Reg. No. 29,027

33,329

July 23, 2001  
GREENBLUM & BERNSTEIN, P.L.C.  
1941 Roland Clarke Place  
Reston, VA 20191  
(703) 716-1191

MARKED-UP COPY OF AMENDED CLAIMS

5. (Amended) The method according to claim 1 [any one of claims 1 to 4], in which the gap information is described by using the residue number and/or the number of residues which indicates the location and the length of a gap region existing in the alignment information of more than two sequences, or a numerical data convertible to those numbers by calculation.

6. (Amended) The method according to claim 1 [any one of claims 1 to 4], in which the gap information is described by using a data containing the residue number of another sequence not included in the alignment or a virtual sequence or column number of a standard representation of the alignment information or a numerical data convertible to the data.

7. (Amended) A method of generating a gap information of a new alignment information by calculating on only a gap information generated from an information of one or more alignments by any one of the methods according to claim 1 [claims 1 to 6].

8. (Amended) A method of obtaining a standard representation of an alignment information from both a gap information and a sequence information obtained by the method according to claim 1 [any one of claims 1 to 7].

12. (Amended) A data record including at least its identifier and a gap information and identifiers of sequences in a sequence information generated by the method according to claim 1 [any one of claims 1 to 10], and a media containing one or more of the record.

13. (Amended) The data record according to claim 12, wherein its identifier is generated by [the method according to claim 11] a method of determining a substantially unique identifier of an alignment information, wherein the identifier is determined using and depending on only and all of a data on a gap information and identifiers of sequences in a sequence information.

14. (Amended) A method of storing, searching and/or communicating method of an alignment information which employs at least the data record according to claim 12 [or claim 13].

15. (Amended) The method according to claim 1 [any one of claims 1 to 14] which employs the correspondence information instead of the gap information.

16. (Amended) The method according to claim 1 [any one of claims 1 to 15] which employs eigen-identifier of sequences in the sequence information.



## Claims

1. A description method of an alignment information characterized by the separation of an alignment information on an amino-acid sequence or a nucleic-acid sequence into a sequence information and a gap information expressing correspondence between sequences.

2. A storing method of an alignment information characterized by the separation of an alignment information on an amino-acid sequence or a nucleic-acid sequence into a sequence information and a gap information expressing correspondence between sequences, and wherein each information is stored and/or searched in one or more storage devices.

3. The method according to claim 2 which stores at least the gap information in one or more storage device.

4. The method according to claim 2 which stores only the gap information.

5. The method according to any one of claims 1 to 4, in which the gap information is described by using the residue number and/or the number of residues which indicates the location and the length of a gap region existing in the alignment information of more than two sequences, or a numerical data convertible to those numbers by calculation.

6. The method according to any one of claims 1 to 4, in which the gap information is described by using a data containing the residue number of another sequence not included in the alignment or a virtual sequence or column number of a standard representation





## DESCRIPTION

Describing and Storing Method of  
Alignment Information

## Field of Invention

The invention relates to describing and storing method of an alignment information with less data than its standard representation, where an alignment information is arranged in such a way as to correspond as many as possible in similar amino-acid residues among multiple amino-acid sequences or nucleic-acid residues among multiple nucleic-acid sequences.

## Background Art

Information on sequences of nucleic acids bearing genetic information is translated to amino-acid sequences. Although functions and structure of a protein are determined by the sequence of 20 kinds of amino-acid residues, it is difficult to predict its functions and structure only from its amino-acid sequence. Thus, it requires a great deal of experimental effort for obtaining those kinds of knowledge. Recently as a result of genome analysis, many amino-acid sequences in various organisms have been determined without isolating the proteins experimentally. In spite of increased amount of sequence information, the functional roles of about half of the amino-acid sequences are still unknown. Therefore, informational techniques for studying the functions and structures of amino-acid sequences become more important.

Generally, it is possible to predict the functions and structure of a protein based on the empirical rule that the higher the similarity between the amino-acid sequences of two proteins, the higher the probability of having the same functions and similar structures. Moreover, it is also possible to estimate dendrograms of species depending on the similarities of the proteins extracted from those organisms. In those cases, it is important to create an "alignment" in which similar amino-acid residues among multiple sequences are corresponded as many as possible. The term "alignment" means processes of finding the correspondence of similar amino acids among the multiple sequences and also means a chart of the aligned sequences representing the correspondence as a result. In most cases, the correspondence of residues is evaluated by using score functions that reflect the similarity or distance of the residues. (The following explains the alignment of the amino acids in more details, where the term "alignment" used herein includes the alignment of the nucleic acids). Alignment information is expressed as aligned letters to represent the correspondence of amino-acid residues which are lined vertically (as a review, see, 「3. Compare and Arrange Letters by Shigeaki Mitaku and Minoru Kanehisa, Baifukan, 1995」).

The alignment is a means to be utilized in science and industry, and studies which require alignment are increasing. Alignments are frequently carried out among the known proteins, for example, examining the relationship between functions and structures of multiple proteins with the same functions in different species, constructing a structure model of a protein based on crystal

structures of similar proteins with known structures instead of crystal analysis. As a result, alignment results are often published on the journals of biochemistry, molecular biology, gene engineering and the like. However, the fact is that each researcher makes an alignment from the sequences as the need arises, and alignment information is often discarded after use. By making a database of alignment information, it is expected to apply the information to other study. Consequently, effectiveness of research will be promoted. Since the alignment technique has been almost standardized, it seems useful to add the alignment information among known proteins and newly known proteins by the genome analysis to the above database.

Although the standard representation of alignment information that expresses the gaps by inserting hyphens into the amino-acid sequence is convenient for a researcher to understand the similarity visually, the data structure of the representation is not appropriate for storing enormous alignment information compactly in a storage device, because it is necessary to record all letters expressing the residues of sequences and gaps. The total data size is at least (the number of residues in each sequence + the number of gaps)  $\times$  number of sequences. The stored data of above form tends to be wasted from the point of information management, and also it leads to storage of redundant information because sequence information itself is usually obtained from the sequence information database. Since alignment information contains information of sequences therein, the redundancy of sequence information becomes terrible in many alignments among which the



and each row includes residues of a sequence (Table 1). Hereinafter in the specification, this representation of the alignment information is called a "standard representation." Amino-acid residues placed in the same column (vertical direction) are related to each other, and a hyphen is inserted when no corresponding residue exists in either sequence. Such a hyphen or a train of hyphens is called "gap region" or simply "gap." "Number of residues in gap" is equal to the length of the gap or the number of the hyphens in the gap. Other regions except gaps are all corresponding.

Table 1

Standard representation of Alignment information of amino-acid sequences

Sequence A --MISLIAALAVD-VIMGRHTWESIVYEQFLPKAQHDLYIA--

Sequence B RSMLSIVAVCQNDVIMGKKTWFSIVY----AKAQHEK FVSPA

The Inventors recognized that the alignment information can be separated into a "sequence information" and a "gap information" which contains information on the residue number at which each gap is inserted and the length of the residues of the gap, or the residue number and the number of residues of other sequences corresponding to the gap region of other sequence, or "correspondence information" which contains information on the residue number and the number of residues of each corresponding region. The sequence information is an array of characters in which each character represents one of 20 kinds of amino acids or 4 kinds of nucleic acids. The sequence



information includes only the information on the sequence itself and does not include any information on correspondence with other sequences. The gap information or the correspondence information is a numerical data expressed by the residue number or the number of residues, both of which are equivalent information and mutually convertible. Consequently, it was found that a conversion of those separated forms of information to standard representation of alignment information was easily achieved by combining the gap information (or the correspondence information) with the sequence information.

Generally the number of gap regions in an alignment is small comparing with the number of amino-acid residues, which is about less than one-tenth. Therefore, it is possible to transmit on communication or to store alignment information effectively with very small amount of data by separating the gap information (or the correspondence information) and the sequence information from the alignment information and by storing or transmitting only the gap information (or the correspondence information). Furthermore, as it is often possible to obtain sequence information from available sequence database, only a gap information (or a correspondence information) can be handled to store or communicate an alignment information. When the sequence information is not readily available, the gap information together with the sequence information can be stored or transmitted. Since a sequence information is included in a standard representation of an alignment information, if this form of alignment information and sequence information is stored or transmitted, the sequence information will





alignment information from both the gap information and the sequence information generated by any one of the aforementioned methods (1) through (7);

(9) A communication method of an alignment information characterized by the separation of an alignment information on an amino-acid sequence or a nucleic-acid sequence to a sequence information and a gap information expressing correspondence between sequences, and by the communication of the gap information at least out of these information;

(10) The aforementioned method (9) in which redundancy of sequence information is removed prior to communication and necessary minimum of the sequence information is transferred;

(11) A method of determining a substantially unique identifier of alignment information, wherein the identifier is determined using and depending on only and all of the data on the gap information and identifiers of sequences in the sequence information;

(12) A data record including at least its identifier and the gap information and the identifiers of sequences in the sequence information generated by any one of the aforementioned methods (1) through (10), and a media containing one or more of the record;

(13) The aforementioned data record (12), wherein its identifier is generated by the aforementioned method (11);

(14) A method of storing, searching, and/or communicating an alignment information which employs at least the aforementioned data record (12) or (13);

(15) Any one of the aforementioned methods (1) through (14) which employs the correspondence information instead of the gap





scope of present invention is not limited to the following embodiments or to any details of descriptions. Moreover, although the following explanations refers only to embodiments of separating alignment information into the sequence information and the gap information, it should not be interpreted that the method of the present invention is limited to those using the gap information, because the "gap information" and the "correspondence information" are equivalent and convertible mutually.

In general, the term "data record" used herein means a unit of storage including one or more representations of an information, an object of search (the data record being a file, a record in a database, or the like). The data records are usually stored in a media accessible by computers. The term "database" used herein means an apparatus comprising a storage means of one or more data records and an access means to any of the data records stored in media or storage devices.

The term "eigen-identifier of a sequence" means a substantially unique identifier of a sequence in a system wherein identifiers of the same sequences are the same and those of different sequences are substantially different (Japanese Patent Application No.(Hei)11-227438/1999). More specifically, the aforementioned eigen-identifier is computed by using a function whose domain is a set of various representations of the sequences and range is a set of their identifiers. For example, the function takes sequential characters expressing each residue in the sequence as its input argument. Then it converts them to a standard representation of the sequence (standardization process), because

the same sequence can be represented in different notations. And the function transforms the representation to a sequence of bits of fixed length using one or more collision intractable hash functions, and then transforms the sequence of bits to a sequence of characters by converting each 4 or 5 bits to a single character. A preferred example of the hash functions includes SHA1, which has properties that takes arbitrary length of input data and gives 160-bit length of output value, and that SHA1 is virtually guaranteed to produce a different value from a different input data. Thus, regardless of the notation style of the representation of sequence information, the generated eigen-identifier becomes unique to the sequence information.

The term "sequence" used herein include both an amino-acid sequence and a nucleic-acid sequence unless otherwise specifically mentioned. The term "alignment information" used herein should be understood in the broadest sense including alignment results published on the journals and alignment results described by standard representations obtained from the standard methods, alignment results described by other expression other than the standard representation, alignment results produced as intermediate data in processes of various analyses, and partial informations of alignment results. The term "store" used herein includes the use of stored data records, as identity means of the data records in databases, existence means for determining whether or not the data record exists in databases, associating means among the data records by using their identifiers, search means for the data records in databases, copying means of any one of the data records to other



databases, duplication means of all of the data records in databases, backup means of the data records by duplication means, access means of the data records in databases, recovery means from storage created by backup means, integration means of more than two data records to a single data record and the like. The term "communication" used herein means copying of a data from a location (sender) to another location (receiver) in a system, where the data is stored in a media or a storage device at each location. Thus, a sender and a receiver may be different memory addresses of the same computer, or may be different computers joined by networks.

By removing the sequence information from the alignment information on the two sequences A and B as shown in Table 1, the information shown in Table 2 is obtained. The number indicates the residue number, and the kind of amino acid corresponding to each residue number can be deducted from the amino-acid sequence information. Therefore, in the methods of present invention, the location and the length of the gap region expressed by hyphens shown on Table 2 are described and stored in other forms.

Table 2

A	-	-	1	2	3	4	5	6	7	8	9	10	11	-	12	13	14	15	16	17	18
B	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
A	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	-	
B	22	23	24	25	26	27	-	-	-	-	28	29	30	31	32	33	34	35	36	37	

0000012-410804

The method of the present invention is characterized by storing an alignment information using a small amount of numerical data without the sequence information shown in Table 2, and as a means for that, it employs a "gap information" to represent correspondences between the sequences. In the following, a practical method of extracting a gap information from the alignment information is explained. However, the kinds or the representation of a gap information are not limited to the following method. As a gap information, any information which represents correspondence between sequences may be utilized. Information on the number of residues on each sequence is assumed to be included in the sequence information in the following explanation.

#### <Representation of Gap Information>

Alignment indicates relative relationship between multiple sequences. The location and the length of a gap region in an aligned sequence of an alignment can be represented by the residue number indicating the location of the region and the number of hyphens contained in the region. Representation methods are not limited, therefore any method may be employed. Some typical embodiments are given in the following, however, the representation methods are not limited to these embodiments. Methods are divided roughly into following ①, ②, and ③. Method ① is a representation based on the residue number of one selected sequence among multiple sequences included in the alignment. Usually, the first sequence in the alignment is selected. Method ② is represented by using the column number of the alignment described by the standard

representation, or by using the residue number of a virtual sequence which is not a real sequence included in the alignment but has no gap and the same length as column length of the alignment, thus its residue number serves as the same as the column number of the alignment. In method ③, gap information is represented by placing the number of residues of the gap region and the residue region alternately for each sequence of the alignment. In each method, various modifications and alternations are possible. In case when the gap information is expressed by the number of residues only, it is desirable to describe gap regions in the increasing order of residue number.

Assume sequence A as the selected sequence, gap information of the alignment on Table 2 is represented as [2, 11, 1, 13, -4, 9, 1] by arranging each residue number alternately (residue number of gap region first in this example) of the gap region and the corresponding region (method ①a). In order to show where the gap region exists in which sequence, the residue number of the gap in sequence B is expressed as a negative number by putting a minus sign. This representation means that from the left end gap of 2 residues in sequence A, corresponding region of 11 residues, gap of 1 residue in sequence A, corresponding region of 13 residues, gap of 4 residues in sequence B, corresponding region of 9 residues, and gap of 1 residue in sequence A at the end are lined up.

Alignment shown in Table 2 may also be represented as [0, 2, 11, 1, 24, -4, 37, 1] by using residue number and the number of residues when the residue number is placed first (method ① b). This representation expresses each gap region by the residue



Table 3

A	-	-	1	2	3	4	5	6	7	8	9	10	11	-	12	13	14	15	16	17	18	19
B	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
C	-	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
A	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	-				
B	24	25	26	27	-	-	-	-	28	29	30	31	32	33	34	35	36	37				
C	23	24	25	-	-	-	26	27	28	29	30	31	32	33	34	35	36	-				

The number of columns in a standard representation of an alignment information is usually larger than the number of the residues of the longest sequence due to gap. For the representation of the alignment in Table 3, a serial number is given to the columns as shown in Table 4 (rows of column numbers are indicated as R)

Table 4

R	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
A	-	-	1	2	3	4	5	6	7	8	9	10	11	-	12	13	14	15	16	17	18	19
B	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
C	-	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

R	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
A	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	-
B	23	24	25	26	27	-	-	-	-	28	29	30	31	32	33	34	35	36	37
C	22	23	24	25	-	-	-	26	27	28	29	30	31	32	33	34	35	36	-

#### Method ②

Based upon the column numbers (R as in Table4), the location of residues or gaps in each sequence are expressed by method ②. When the first and the last column numbers of the residues from the N-terminal are described by this method, the alignment in table 3 is represented as follows:

Sequence A    3, 13, 15, 40

Sequence B    1, 27, 32, 41

Sequence C    2, 26, 30, 40

In sequence A, residues exist between the column numbers 3 and 13, 15 and 40. For sequence B, residues exist between the column numbers 1 and 27, 32 and 41. For sequence C, residues exist between the column numbers 2 and 26, 30 and 40.

#### Method ③

For each sequence, it is possible to represent the gap information by indicating the number of residues of each gap and corresponding residues alternately without using any specifically selected sequence or column number as basis (method ③). According to this method, when the number of residues of gap region is shown first, Table 3 alignment can be described as follow:

Sequence A    2, 11, 1, 26, 1

Sequence B    0, 27, 4, 10

Sequence C    1, 25, 3, 11, 1

In the gap information of sequence A, 2, 1, and 1 show the number of residues in each gap, and continuous 11 and 26 amino-acid residues are placed between the gaps. In sequence B, after the 27 residues from the N-terminal, 10 residues exist across the gap of 4 residues. In sequence C, 25 residues exist after the gap of 1 residue of N-terminal, 11 residues exist after the gap of 3 residues, and gap of 1 residue at the end exist. It is also possible to begin with the number of corresponding residues first.

#### <Conversion to Standard Representation of Alignment Information>

A gap information represented by any method can be converted to a standard representation of an alignment information. First, it is necessary to calculate the number of columns required to enumerate all residues including the gap and to prepare columns. Following the gap information, the alignment information in Table 3 will be obtained when the residue numbers or hyphens are corresponded in each column. Moreover, the standard representation of alignment information will be reproduced when amino-acid residues

are placed into the corresponding residue number of each sequence.

The alignment information shown in Table 3 will be reproduced as follows. From the gap information in Table 2 according to alignment by method ② (sequence A 「3, 13, 15, 40」, sequence B 「1, 27, 32, 41」, sequence C 「2, 26, 30, 40」), necessary number of columns is determined as equal to the largest column number, i.e., 41. After 41 columns are prepared, for sequence A all 37 residues are placed in each column from the N-terminal in columns from numbers 3 to 13 and 15 to 40. For sequence B, all 37 residues are placed in each column from the N-terminal in columns from numbers 1 to 27 and 32 to 41, and for sequence C, all 36 residues are placed in each column from number 2 to 26 and 30 to 40.

From the gap information shown in Table 3 according to alignment by method ③ (sequence A 「2, 11, 1, 26, 1」, sequence B 「0, 27, 4, 10」, sequence C 「1, 25, 3, 11, 1」, necessary number of columns is calculated as  $2+11+1+26+1=41$  by adding all residue numbers of sequence A. The same number is obtained by calculation based on sequence B ( $27+4+10=41$ ) or sequence C ( $1+25+3+11+1=41$ ). In these columns, for sequence A, 11 residues after the gap of 2 residues, 26 residues after the gap of 1 residue, and a gap of 1 residue are lined in order from the N-terminal. For sequence B, 27 residues are lined in order from the N-terminal, and 10 residues after the gap of 4 residues are lined. The same procedure is applied to sequence C.

Generally, the column length of an alignment often changes in a standard representation of an alignment by modification such as deletion of one or more of the sequences contained or addition



of other sequences. The change of the column length of an alignment is caused dependently on the change of insertion mode of gaps, whilst independently on the sequence information. The calculation method of the present invention is characterized in that the gap information of a modified alignment is easily converted from a gap information of one or more original alignments, with no sequence information required.

#### <Extraction of Sequence from Alignment Information>

The following is a procedure of extracting alignment information of A and C according to the standard representation from the alignment information shown in Table 3 by removing one part of sequence, for example, sequence B. From the gap information set out in Table 3 according to alignment by method ② (Sequence A 「3, 13, 15, 40」, Sequence B 「1, 27, 32, 41」, and Sequence C 「2, 26, 30, 40」), the information on sequence A and sequence C (「3, 13, 15, 40」, 「2, 26, 30, 40」) is extracted. In the columns from 1 to 41, column numbers which correspond to the gap in both sequences (in this case 1 and 41) are searched operationally. In the example, the smallest number of the gap information is 2 in sequence C, therefore, decrement of all the numbers is made so that the smallest becomes 1 which is the starting end of alignment. Gap informations of sequence A and C are 「2, 12, 14, 39」 and 「1, 25, 29, 39」, respectively. The gap region between column number 12 and 14 of sequence A does not overlap with the gap region between column number 25 and 29 of sequence C. As a result, the number of columns becomes 39, and these column numbers are moved to the

left end (small column number side). The alignment obtained from the calculated gap information is shown in Table 5.

Table 5

R	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
22																						
A	-	1	2	3	4	5	6	7	8	9	10	11	-	12	13	14	15	16	17	18	19	20
C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
22																						
R	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39					
A	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37					
C	23	24	25	-	-	-	26	27	28	29	30	31	32	33	34	35	36					

As for the gap information for the alignment shown in Table 3 by method ③ (sequence A 「2, 11, 26, 1」, sequence B 「0, 27, 4, 10」, sequence C 「1, 25, 3, 11, 1」, after the conversion to gap information by method ②, extraction of sequences will be easily conducted by the calculation of gap information as explained above.

For sequence A, an example of conversion from the gap information by method ③ to that by method ② will be shown below. The gap information obtained by method ③ 「2, 11, 26, 1」 describes the number of residues of gap regions and residue regions alternately, and there are two places where the residues exist. The first and the last column numbers of each region can be calculated as  $2+1=3$ ,  $2+11=13$  and  $2+11+1+1=15$ ,  $2+11+1+26=40$ , and accordingly, it is

possible to convert the information to the gap information by method ②, i.e., 「3, 13, 15, 40」. When the same procedure is applied to sequence B, the gap information by method ③ 「0, 27, 4, 10」 can be converted to the gap information by method ② 「1, 27, 32, 41」 by the operation of  $0+1=1$ ,  $0+27=27$ ,  $0+27+4+1=32$ ,  $1+27+4+10=41$ .

#### <Integration of Multiple Alignment Information>

According to the methods of present invention, more than two alignment results where common sequences exist, such as shown in Table 6, can be easily integrated by the calculation of gap information.

0959312-110801

Table 6

## Alignment 1

A - - 1 2 3 4 5 6 7 8 9 10 11 - 12 13 14 15 16 17 18 19

20

B 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

23

C - 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

22

A 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 -

B 24 25 26 27 - - - 28 29 30 31 32 33 34 35 36 37

C 23 24 25 - - - 26 27 28 29 30 31 32 33 34 35 36 -

## Alignment 2

A - - - 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

20

D 1 2 3 4 5 6 7 8 9 10 11 12 13 - 14 15 16 17 18 19 20 21

22

A 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 -

D 23 24 25 - - 26 27 28 29 30 31 32 33 34 35 36 37 -

The gap information in alignment 1 by method ② is 「3, 13, 15, 40」 for sequence A, 「1, 27, 32, 41」 for sequence B, and 「2, 26, 30, 40」 for sequence C, and the gap information in alignment 2 is 「4, 40」 for sequence A and 「1, 13, 15, 26, 29, 40」 for sequence

D. From the common gap information of sequence A, it can be understood that placements of new gap of 1 residue to the N-terminal for alignment 1 and new gap of 1 residue between column numbers 14 and 15 for alignment 2 are required.

Therefore, in all sequences contained in alignment 1, as a result of adding 1 to each column number due to the N-terminal gap, the gap information for sequence A is 「4, 14, 16, 41」, for sequence B 「2, 28, 33, 42」, and for sequence C 「3, 27, 31, 41」. For alignment 2, due to the introduction of the new gap between the column numbers of 14 and 15, the gap information for sequence A is 「4, 14, 16, 41」 and for sequence D 「1, 13, 16, 27, 30, 41」. Thus, in both alignments, identical gap informations on sequence A are obtained. If the above information is converted to a standard representation according to the above mentioned procedures, integrated alignments shown in Table 7 are obtained. The required number of columns in the integrated alignments is 42 based on the largest column number of sequence B.





the gap information and identifiers of sequences in the sequence information.

Sequence information is not necessarily stored in the files, tables or database containing gap information. Although it is preferable to store a sequence information in the same media as that stores a gap information, a sequence information may be stored in other medias available for searching. When the sequence information stored in other database is available, only the gap information may be stored in the database. It is also possible to store some sequence informations frequently used together with the gap information, and to utilize an outside database for other sequence informations. Furthermore, a sequence information in the database may include species and organisms and subtypes in addition to the sequence identifier, the name of proteins, the number of amino acid residues, amino-acid sequence, and the like. The information may be administered separately by the tables of a relational database.

#### <Communication Method>

It is possible to communicate an alignment information effectively by transmitting the alignment information from a sender to a receiver in a separated form of a sequence information and a gap information. First, according to the methods described above, an alignment information is separated to a sequence information and a gap information. A sequence identifier of an amino-acid sequence corresponding to the gap information is added to the gap information, which is transmitted from the sender to the receiver.





sequence information by other methods specifically explained above or those not disclosed in the specification.

#### Example 1

Alignment information on 4 amino-acid sequences was divided to the gap information and the sequence information as shown in Table 8, and the gap information was stored in the database. In the Table, each amino-acid sequence was specified by giving the sequence identifier. The term "ID" herein means identifier. In the gap information, sequence ID=000001 represents the selected sequence as a basic sequence, and the gap information of sequence ID=000002 through 000004 are represented relative to the selected sequence.

Table 8

#### (Gap Information)

Sequence ID	Gap Information
000001	3, 11, 1, 26, 1
000002	1, 27, 4, 10, 0
000003	2, 25, 3, 11, 1
000004	0, 13, 2, 12, 2, 12, 1

#### (Sequence Information)

Sequence ID; Protein Name; Number of amino acid residue; and  
Amino-acid sequence

000001	xxxxxxx	37
--------	---------	----

09691-1001

The property information in the columns of alignment (X) in Table 9 is integrated with alignment (Y) in Table 10 and marked.

Alignment information (X)

— \* — \* — \* — — — # — # \* \* \* — — # \* — \* \* \* — — — — \* \* \* — — — —

--MISLIAALAVD-VIMGRHTWESIVYEQFLPKAQHDLYIA-

RSMLSIVAVCQNDVIMGKKTWFSIVY----AKAQHEK FVSPA

Alignment information (Y)

-RSMLSIVAVCON---DAVIMGKKTWFSIVYAKAQHEK FVSPA

A-SVVSLAAVCRNNKPEAVLMMKKSWFSLLYAKAQHEKFVSPV

In Table 9, the positions in which amino-acid sequences are the same in the correspondence of columns in sequence A and sequence B are indicated as \*. Also the positions of amino acids which are functionally important are marked as #. According to the same procedure of separating the alignment information into the sequence information and the gap information as shown in Table 8, the column property information is separated to the property kind information and the column location information as shown in Table 11 by regarding the [-] as gap in the column property information on Table 9. In this case, the column location information is the same as the gap information expressed by method ③.

Table 11

Property kind information	***##*****
Column location information	2,1,1,1,2,1,4,1,1,4,2,2,1,4,5,4,7

Table 12 shows the results of calculated column location information in alignment (Y) shown in Table 10 from the column location information shown in Table 11, the gap information on sequence B shown in Table 9, and the gap information on sequence B shown in Table 10.

Table 12

Column location information 3,1,1,1,2,1,7,1,1,4,2,2,1,4,1,4,7

Table 13 shows the column property information on alignment (Y) obtained from the column location information on Table 12 and the property kind information on Table 11. As it is clear from comparison of Table 9 with Table 13, the column property information is corresponded on sequence B which is common to alignment (X) and alignment (Y).

Table 13

Alignment information (Y) which is mapped so as to correspond with the column information of alignment information (Y) on sequence B.

Column property information

[illegible]

Sequence B

-RSMLSIVAVCON---DAVIMGKKTWGSIVYAKAQHEK FVSPA

Sequence C

A-SVVSLAAVCRNNKPEAVLMMKKSWFSLLYAKAQHEK FVSPV

### Example 3

It is demonstrated that representation of the alignment information shown in Table 14 is separated to the sequence

information (Table 15) and the gap information. The gap information is stored as a data record (Table 16) containing its identifier, the gap information, identifiers of sequences, date and the like. The identifiers of sequences are eigen-identifiers, and the identifier of the data record is determined using and depending on only and all of the data on the gap information and identifiers of the sequences. The data in Table 14 to 16 is written by XML (extensible markup language).

Table 14

Sequence P

-RSMLSIVAVCQN---DAVIMGKKTWFSIVYAKAQHEK FVSPA

Sequence Q

A-SVVSLLAAVCRNNKPEAVLMMKKSWFSLLYAKAQHEK FVSPV

In Table 15, sequence information of sequence P and Q is shown. Each sequence is tagged by "<sequence>" and "</sequence>" at its head and tail. The start-tag at the head indicates the start of the sequence and the end-tag at the tail indicates the end of the sequence. In the start-tag <sequence>, "ed=" is an attribute that indicates the eigen-identifier of the sequence between the tags. Thus, the eigen- identifier of sequence P is:

"SA16rxgd7d4xxgmjcuaf8v3f6crqu8p9bck."

The eigen- identifier of sequence Q is:

"SA1jlr9pr0f9xcc00p57xke0kdijp8jvrh4."

Table 15

```
<sequence ed="SA16rxgd7d4xxgmjcuaf8v3f6crqu8p9bck">
RSMLSIVAVCQNDVIMGKKTWFSIVYAKAQHEK FVSPA
</sequence>
```

```
<sequence ed="SA1jl9pr0f9xcc00p57xke0kdijp8jvrh4">
ASVVSLAAVCRNNKPEAVLMMKKSWFSLLYAKAQHEK FVSPV
</sequence>
```

Table 16

```
<record ed=" AL17vr5emniqtmnj36eir4r0vbny5ymlnj9"
date="19990507">
<reference gap="1,12,3,27" order="1">
SA16rxgd7d4xxgmjcuaf8v3f6crqu8p9bck
</reference>
<reference gap="0,1,1,41" order="2">
SA1jl9pr0f9xcc00p57xke0kdijp8jvrh4
</reference>
</record>
```

In Table 16, the data record including gap information is shown. The data record starts from <record> tag, and ends at </record> tag. Thus, the data record is consisted of whole letters in Table 16. In the data record, the identifier of each sequence is placed between <reference> and </reference>. In each

<reference> tag, gap information (method ③) is expressed by "gap=".  
" The "order=" in each <reference> tag indicates the dictionary order of the identifier in the data record.

The "ed=" in the <record> tag indicates the identifier of the record. The identifier is included in the data record as shown in Table 16. The following explains that the identifier is substantially unique identifier of the data record, where the date information does not effect on its uniqueness. In order to generate the identifier based on the gap information and identifiers of the sequences, each sequence identifier is joined with the letters in its "gap=," and then they are joined in the order of dictionary. Thus after above procedures, it become one string of letters as shown in Table 17. The string is unique to the gap information and the identifiers of sequences, because the string is created from them. The string is then transformed by SHA-1 to 160-bit substantially unique data. Each 5 bits of the 160-bit data converted to one of "0, 1, 2, 3, 4, 5, 6, 7, 8, 9, a, b, c, d, e, f, g, h, i, j, k, x, m, n, y, p, q, r, s, t, u, v," yielding 32-letter unique word. Finally, the substantially unique identifier is created by the word which is joined after "AL1" meaning the identifier is unique to the alignment information. Since the identifier is unique to the alignment information, it is favorable to use it as that of the data record.

Table 17

"SA16rxgd7d4xxgmjcuaf8v3f6crqu8p9bck1,12,3,27  
SA1j1r9pr0f9xcc00p57xke0kdi jp8jvrh40,1,1,41"





and it is possible to store a huge amount of alignment informations and prepare database more effectively for the reuse of those alignment informations.

Furthermore, since a sequence information and a gap information are administered separately, it will be easy to preserve the consistency and maintenance of database. Particularly in a relational database, possibility of the use of database will increase because the data is treated in more canonical conditions. Moreover, when the alignment information is transmitted according to the methods of present invention, receivers do not have to receive the sequence information already at hand, and accordingly, communication efficiency will improve and the overlap of the sequence information on the receiver side will be avoided. Particularly, it is effective when a huge amount of alignment informations is transmitted, duplication of database through communication is made, and the alignment information in the client-server system and communication between applications is exchanged.

## Declaration and Power of Attorney for Utility or Design Patent Application

## 特許出願宣言書

## Japanese Language Declaration

私は、下欄に氏名を記載した発明者として、以下のとおり宣言する：

私の住所、郵便の宛先および国籍は、下欄に氏名に続いて記載したとおりであり、

名称の発明に関し、請求の範囲に記載した特許を求める主題の本来の、最初にして唯一の発明者である（一人の氏名のみが下欄に記載されている場合）か、もしくは本来の、最初にして共同の発明者である（複数の氏名が下欄に記載されている場合）と信じ、

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

上記発明の明細書（下記の欄で X 印がついていない場合は、本書に添付）は、

☐ \_\_\_\_\_年\_\_\_\_月\_\_\_\_日に提出され、米国出願番号

\_\_\_\_\_とし、（該当する場合）

\_\_\_\_\_年\_\_\_\_月\_\_\_\_日に訂正されました。又は、

特許協定条約国際出願番号\_\_\_\_\_とし、

（該当する場合）\_\_\_\_\_年\_\_\_\_月\_\_\_\_日に訂正されました。

私は、前記のとおり補正した請求の範囲を含む前記明細書の内容を検討し、理解したことを陳述する。

私は、連邦規則法典第 37 編第 1 条 56 項に定義されているとおり、特許資格の有無について重要な情報を開示すべき義務があることを認めます。

私は、合衆国法典第 35 部第 119 条（a-d）項又は第 365 条（b）項に基づく、下記の外国外特許出願又は発明者証出願、或いは第 365 条（a）項に基づく、少なくとも米国以外の 1 カ国を指名した PCT 国際出願の外国優先権を主張し、更に優先権の主張に係る基礎出願の出願日前の出願日を有する外国特許出願、又は発明者証出願あるいは PCT 国際出願を以下に“なし”の箱に印をつけることにより明記する：

## Prior foreign applications

先の外国外出願

11/15189 Japan 25/Jan/99  
(Number) (Country) (Day/Month/Year Filed)  
(番号) (国名) (出願の年月日)

(Number) (Country) (Day/Month/Year Filed)  
(番号) (国名) (出願の年月日)

☐ その他の外国外特許出願番号は別紙の追補優先権欄にて記載する。

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name:

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

**Describing And Storing Method Of Alignment Information**

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

the specification of which is attached hereto unless the following box is checked:

☒ was filed on 25/Jan/00 as United States Application Number 09/869,312 and was amended on 23/Jul/01 (if applicable) or,

PCT International Application Number  
PCT/JP00/00355 and was amended on  
\_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority under Title 35, United States Code §119(a-d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT international application which designated at least one country other than the United States, listed below. I have also identified below, by checking the "No" box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed:

Priority claimed

優先権の主張

☒ Yes ☐ No  
あり なし

☐ Yes ☐ No  
あり なし

☐ Additional foreign application numbers are listed a supplemental priority sheet attached hereto.

# Japanese Language Utility or Design Patent Application Declaration

私は、合衆国法典第 35 部第 119 条 (e) 項に基づく、下記の合衆国特許出願の利益を主張する。

(Application No.)  
(出願番号)

(Application No.)  
(出願番号)

(Application No.)  
(出願番号)

☐ その他の合衆国特許出願番号は別紙の追補優先権欄に記載する。

私は、合衆国法典第 35 部第 120 条に基づく下記の合衆国特許出願、又は第 365 条 (c) 項に基づく合衆国を指名した PCT 国際出願の利益を主張し、本願の請求の範囲各項に記載の主題が合衆国法典第 35 部第 112 条第 1 項規定の様態で、先の合衆国特許出願又は PCT 国際出願に開示されていない限度において、先の出願の出願日と本願の国内出願日又は PCT 国際出願日の間に有効となった連邦規則法典第 37 部第 1 章第 56 条に記載の特許要件に所要の情報を開示すべき義務を有することを認める。

(Application No.)  
(出願番号)

(Day/Month/Year Filed)  
(出願の年月日)

(Application No.)  
(出願番号)

(Day/Month/Year Filed)  
(出願の年月日)

☐ その他の合衆国又は国際特許出願番号は別紙の追補優先権欄に記載する。

私は、ここに自己の知識に基づいて行った陳述が全て真実であり、自己の有する情報および信ずるところに従って行った陳述が真実であると信じ、さらに故意に虚偽の陳述等を行った場合、合衆国法典第 18 部第 1001 条により、罰金もしくは禁に処せられるか、またはこれらの刑が併科され、またかかる故意による虚偽による陳述が本願ないし本願に対して付与される特許の有効性を損なうことがあることを認識して、以上の陳述を行ったことを宣言する。

私、下記署名者は、ここに記載の米国弁護士または代理人に本出願に関し特許商標庁にて取られるいかなる行為に関して、同米国外弁護士又は代理人が私に直接連絡なしに私の外国弁護士或いは法人代表者からの指示を受け取り、それに従うようここに委任する。この指示を出す者が変更の場合には、ここに記載の米国弁護士又は代理人にその旨通知される。

I hereby claim the benefit under Title 35, United States Code §119 (e) of any United States provisional application(s) listed below.

(Day/Month/Year Filed)  
(出願の年月日)

(Day/Month/Year Filed)  
(出願の年月日)

(Day/Month/Year Filed)  
(出願の年月日)

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s), or §365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

(現況) (Status)  
(特許済み、係属中 放棄済み) (patented, pending, abandoned)

(現況) (Status)  
(特許済み、係属中 放棄済み) (patented, pending, abandoned)

☐ Additional U.S. or international application numbers are listed on a supplemental priority sheet attached hereto.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

The undersigned hereby authorizes the U.S. attorney or agent named herein to accept and follow instructions from either his foreign patent agent or corporate representative, if any, as to any action to be taken in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorney or agent and the undersigned. In the event of a change in the persons from whom instructions may be taken, the U.S. attorney or agent named herein will be so notified by the undersigned.

## Japanese Language Utility or Design Patent Application Declaration

委任状: 私は、下記発明者として、下記に明記された顧客番号を伴う以下の弁護士又は、代理人をここに選任し、本願の手続きを遂行すること並びにこれに関する一切の行為を特許商標庁に対して行うことを委任する。そして全ての通信はこの顧客番号宛に送られる。

顧客番号 7055

現在委任された弁護士は下記の通りである。

Neil F. Greenblum  
Bruce H. Bernstein  
James L. Rowland  
Arnold Turk

Reg. No. 28,394  
Reg. No. 29,027  
Reg. No. 32,674  
Reg. No. 33,094

POWER OF ATTORNEY: As a named inventor, I hereby appoint the attorney(s) and/or agent(s) associated with the Customer Number provided below to prosecute this application and transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

CUSTOMER NUMBER 7055

The appointed attorneys presently include:

Stephen M. Roylance  
William E. Lyddane  
William Pieprz  
Leslie J. Paperner

Reg. No. 31,296  
Reg. No. 41,568  
Reg. No. 33,630  
Reg. No. 33,329

Address: **GREENBLUM & BERNSTEIN, P.L.C.**  
1941 Roland Clarke Place  
Reston, VA 20191

直接電話連絡先:

Direct Telephone Calls to:

**GREENBLUM & BERNSTEIN, P.L.C.**  
(703) 716-1191

唯一のまたは第一の発明者の氏名	1-00	Full name of sole or first inventor	Tetsuro TOYODA
同発明者の署名	日付	Inventor's signature	Date October 12, 2001
住所		Residence	Tokyo, Japan JPX
国籍		Citizenship	Japan
郵便の宛先		Post Office Address	1-65-5, Mitsufuji, Musashimurayama-shi, Tokyo 208-0021, Japan
第二の共同発明者の氏名 (該当する場合)	2-00	Full name of second joint inventor, if any	Akiko ITAI
同第二共同発明者の署名	日付	Second inventor's signature	Date October 12, 2001
住所		Residence	Tokyo, Japan JPX
国籍		Citizenship	Japan
郵便の宛先		Post Office Address	5-16-6, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

(第三またはそれ以降の共同発明者に対しても同様な情報および署名を提供すること。)

(Supply similar information and signature for third and subsequent joint inventors.)